

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - protein search, using bw mode1.
Run on: October 29, 2005, 00:51:44 ; Search time 69 Seconds
Perfect score: 147 (without alignments)
Title: US-09-980-263-1
Sequence: 1 GMQGPAGSGWEEGSSGSPGVTPFLFSP 26

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: GeneseqP1980s:*

2: GeneseqP1990s:*

3: GeneseqP2000s:*

4: GeneseqP2001s:*

5: GeneseqP2002s:*

6: GeneseqP2003as:*

7: GeneseqP2003bs:*

8: GeneseqP2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	147	100.0	26	AAG62388	Aag62388 Alternative
2	147	100.0	26	AAB48914	Aab48914 Human ace
3	147	100.0	26	AAB5032	Aab5032 Acetylch
4	147	100.0	26	AAU98024	Aau98024 Human rea
5	147	100.0	31	AAW74588	Aaw74588 Amino aci
6	147	100.0	31	AAW68146	Aaw68146 Human ACh
7	147	100.0	37	AAB50035	Aab50035 Acetylch
8	147	100.0	53	AAB50036	Aab50036 Acetylch
9	147	100.0	53	ABG3131	Abg3131 GFP-fused
10	147	100.0	600	AAW8797	Aaw8797 Human ace
11	61.5	42.9	145	ADC33366	Adc33366 Human nov
12	61.5	41.8	575	ABB11475	Abb11475 Human R31
13	61	41.5	98	AAB51865	Aab51865 Human sec
14	60	40.8	54	AAR37744	Aar37744 Collagen-
15	60	40.8	54	AAR93255	Aar93255 Collagen-
16	60	40.8	334	ADL91154	Adl91154 Mouse fic
17	60	40.8	334	ADL91152	Adl91152 Human fic
18	60	40.8	633	AAR37746	Aar37746 Collagen-
19	60	40.8	633	AAR93257	Aar93257 Collagen-
20	60	40.8	633	AASW7655	Aasw7655 Collagen-
21	60	40.8	1065	AAR37745	Aar37745 Collagen-
22	60	40.8	1065	AAR93256	Aar93256 Collagen-
23	60	40.8	1065	AAW57654	Aaw57654 Collagen-
24	60	40.8	1466	AAE02534	Aae02534 Bovine al
25	60	40.8	1466	AAE02533	Aae02533 Bovine al

RESULT 1
AAG62388
ID AAG62388 standard; peptide; 26 AA.
XX
AC AAG62388;
XX
DT 31-AUG-2001 (first entry)
XX
DE Alternatively splice AChE product C-terminus AChE-R.
XX
KW Antisense oligonucleotide; acetylcholine esterase; AChE; dystonia;
cholinergic neurotransmission; progressive neuromuscular disorder;
myasthenia gravis; Eaton-Lambert disease; muscular dystrophy; PTSD;
amyotrophic lateral sclerosis; post-traumatic stress disorder;
multiple sclerosis; post-stroke sclerosis; post-injury muscle damage;
excessive re-innervation.
XX
OS Unidentified.
XX
PN WO00136627-A2.
XX
PD 25-MAY-2001.
XX
PP 16-NOV-2000; 2000WO-IL000763.
XX
XX
PA 16-NOV-1999; 99IL-00132972.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Soreq, H., Seidman, S.,
XX
DR WPI ; 2001-33603/35.
XX
PS Disclosure; Fig 1; 124PP; English.

Sequences AAH44810 - AAH44822 represent antisense oligonucleotides targeting the acetylcholine esterase (AChE) mRNA. AChE is involved in the termination of cholinergic neurotransmission, by hydrolysing the neurotransmitter acetylcholine. Mammalian AChE is encoded by one gene but alternative splicing at its 3' end yields three different mRNA transcripts which encode protein with distinct carboxyl termini. All three proteins are catalytically active. AChE has morphogenic, non-catalytic capacities too. AChE antisense oligonucleotides are used in treating or preventing a progressive neuromuscular disorder. Examples of

CC disorders which are treatable using the antisense oligonucleotides
 CC include myasthenia gravis, Eaton-Lambert disease, muscular dystrophy,
 CC amyotrophic lateral sclerosis, post-traumatic stress disorder (PTSD),
 CC multiple sclerosis, dystonia, post-stroke sclerosis, post-injury muscle
 CC damage, excessive re-innervation and post-exposure to AChE inhibitors.
 CC The present sequence represents the C-terminus of an alternatively

RESULT 4	Db	1	GMQGPAGSGWEGSGSPGVTPLFSP 26
AAU98024	ID	AAU98024	standard; peptide; 26 AA.
XX	XX	AAU98024;	
XX	AC		
DT	27-AUG-2002	(first entry)	
XX			
DB	Human readthrough acetylcholinesterase unique region.		
XX			
KW	Human; acetylcholinesterase; single-chain variable fragment; scFv;		
KW	AChE-S; heavy chain variable region; muscle re-inervation;		
KW	progressive neuromuscular disorder; muscle distortion; readthrough;		
KW	myasthenia gravis; neuromuscular junction abnormality;		
KW	Eaton-Lambert disease; muscular dystrophy; amyotrophic lateral sclerosis;		
KW	ALS; post-traumatic stress disorder; PTSD; multiple sclerosis; Dystonia;		
KW	KW post-stroke sclerosis; post-injury muscle damage; excessive re-innervation;		
XX			
OS	Homo sapiens.		
XX			
PN	WO200246422-A1.		
XX			
PD	13-JUN-2002.		
XX			
PF	22-MAY-2001; 2001WO-11000464.		
XX			
PR	04-DEC-2000; 2000IL-00140071.		
XX			
PA	(YISS) YISSUM RES & DEV CO.		
XX			
PI	Soreq H, Seidman S, Shohami E;		
XX			
PR	WPI; 1998-506377/43.		
XX			
DR	Treatment of injury to central nervous system - by administration of		
XX	PT inhibitor of acetyl-cholinesterase production.		
XX			
PS	Disclosure; Page 62; 88pp; English.		
PS	PT	This is the amino acid sequence of a human acetyl-cholinesterase (AChE) variant used in the method of the invention, where inhibitors of AChE are used to treat injury to the central nervous system (CNS). The AChE	
PS	PT	inhibitor can also be used to facilitate transplantation of neuronal	
PS	PT	cells to the CNS of a patient. The inhibitor following injury to the CNS can be treated with the method include epilepsy, stroke, Huntington's disease, head injury, spinal injury, pain, Parkinson's disease, myelin deficiencies, neuromuscular disorders, neurological pain, amyotrophic lateral sclerosis, Alzheimer's disease, and affective	
PS	PT	disorders of the brain.	
PS	Sequence 31 AA;		
XX			
PS	Nucleic acid sequence coding for a single-chain variable fragment (scFv) antibody that has specific affinity for the synaptic variant of acetylcholinesterase (AChE-S), useful for diagnosing a neuromuscular disorder, e.g. Myasthenia gravis.		
PS	Disclosure; Fig 1; 73pp; English.		
XX	The invention relates to a nucleic acid sequence coding for a single-chain variable fragment (scFv) antibody that has specific affinity for the synaptic variant of acetylcholinesterase (AChE-S), where the scFv antibody consists essentially of a polypeptide comprising the binding portion of the heavy chain variable region of an antibody. Also included are an expression vehicle comprising a nucleic acid sequence coding for a scFv antibody that has specific affinity for the synaptic variant of AChE-S, an scFv antibody specifically recognising and binding to the synaptic variant of AChE-S and a method for the diagnosis of a progressive neuromuscular disorder in a mammal, comprising obtaining a sample from the mammal and detecting intensities expression of at least one of the AChE variants in the sample. The single-chain Fv antibody is useful for diagnosing a progressive neuromuscular disorder which involves any one of muscle distortion, muscle re-innervation and neuromuscular junction (NMJ) abnormalities. The disorder is Myasthenia gravis (preferred), Eaton-Lambert disease, muscular dystrophy, amyotrophic lateral sclerosis (ALS), post-traumatic stress disorder (PTSD), multiple sclerosis, Dystonia, post-stroke sclerosis, post-injury muscle damage; excessive re-innervation, or post-exposure to AChE inhibitors. The present sequence represents the unique region of human acetylcholinesterase encoded by a readthrough splice variant		
XX	Sequence 26 AA;		
XX			
Query Match	100.0%	Score 147; DB 5; Length 26;	
Best Local Similarity	100.0%	Pred. No. 1.4e-10;	
Matches 26;	Conservative	Mismatches 0; Indels 0; Gaps 0;	
XX			
Query	1	GMQGPAGSGWEGSGSPGVTPLFSP 26	
XX			
DB	6	GMQGPAGSGWEGSGSPGVTPLFSP 31	
XX			
RESULT 6			
ID	AAW61146		
XX	AAW61146;		
AC			
XX			
DT	05-OCT-1998	(first entry)	
XX			
DE	Human AChE splice variant E1-4-I4-E5.		
KW	Nuclelease resistant; acetylcholinesterase; human; myasthenia gravis; AChE; AChE-		
XX			

KW Parkinson's disease; Alzheimer's disease; central nervous system; neuromuscular junction; cholinergic signalling; brain.
 KW OS Homo sapiens.
 XX XX WO9826062-A2.
 XX XX 18-JUN-1998.
 PD 12-DEC-1997; 97WO-US023598.
 XX 12-DEC-1996; 96US-0035266P.
 PR 13-FEB-1997; 97US-0037777P.
 PR 02-MAY-1997; 97US-00850347.
 PR 21-JUL-1997; 97US-0053334P.
 XX PA (YISS) YISSUM RES & DEV CO.
 PA (KOHN/) KOHN K I.
 PI Soreq H, Seidman S, Eckstein F, Friedman A, Kaufer D;
 XX DR 1998-348522/30.
 XX PT Synthetic nucleic resistant antisense oligodeoxynucleotides - directed against acetylcholinesterase, useful for treating Parkinson's and Alzheimer's diseases and myasthenia gravis.
 PT XX Disclosure; Fig 12; 89pp; English.
 PS This represents the amino acid sequence of a human acetylcholinesterase (AChE) splice variant. The invention provides sequences shown in AAV41278 to AAV41285 that represent synthetic nucleic resistant antisense oligodeoxynucleotides which are capable of selectively modulating human acetylcholinesterase (AChE) production. These oligonucleotides are targeted to a splice junction in a splice variant of AChE mRNA and are capable of selectively modulating human AChE production in the central nervous system and neuromuscular junction. The invention also provides a method for determining the efficacy of these human AChE specific antisense oligonucleotides. These antisense oligonucleotides can be used to restore balanced cholinergic signalling in the brain, particularly related to learning and memory as well as stress disorders. Parkinson's and Alzheimer's disease. They can also be used to reduce production and therefore deposition of AChE in the neuromuscular junctions of patients with e.g. myasthenia gravis. The oligonucleotides work effectively at low doses while avoiding many of the side effects associated with Tacrine and related cholinergic drugs for Alzheimer's disease and pyridostigmine and related drugs for myasthenia gravis.
 XX SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 12 GMQGPAGSGWGRGSGSPGVPPLFSP 37

SQ Sequence 37 AA;

Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 12 GMQGPAGSGWGRGSGSPGVPPLFSP 37

SQ Sequence 37 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

Query Match 100.0%; Score 147; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWGRGSGSPGVPPLFSP 26
 Db 6 GMQGPAGSGWGRGSGSPGVPPLFSP 31

SQ Sequence 31 AA;

PR 02-SEP-1999; 99IL-00131707.
 (YISS) YISUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Soreq H, Meshorer B, Shlaim E, Shoham S;
 PA XX WPI; 2002-490152/52.
 PI XX DR
 XX PT Evaluating effect of drugs on nervous system by comparing effect of drug
 PT challenge by AChE blocker, AChE activity in brain of test animal following
 PT challenge by AChE blocker and comparing it with control group.
 XX
 XX PS PS Example; Page 52; 114pp; English.
 XX
 XX The present invention relates to a method and system for evaluating an
 CC effect on the nervous system of a test drug. The method comprises
 CC comparing the effect of the drug on acetylcholinesterase (AChE) catalytic
 CC activity or isoenzyme variance in a brain of a test animal following a
 CC challenge by an AChE blocker or a blocker of AChE and muscarinic
 CC receptors M1 and M2 (e.g. Pridostigmine) and comparing this effect with
 CC that of a known agent, preferably a non-selective muscarinic receptor
 CC blocker (e.g. Scopolamine) or a specific M1 receptor blocker (e.g.
 CC Pirenzepine). The method is useful for evaluating an effect on the
 CC nervous system of a test drug, including drugs for the treatment of
 CC anxiety conditions, post-traumatic stress, Alzheimer's disease, muscle
 CC malfunction, neurodegenerative disorders, damage resulting from
 CC exposure to xenobiotics, panic, neuromuscular disorders, Parkinson's
 CC disease, Huntington's chorea, muscle fatigue, multiple chemical
 CC sensitivity, autism, multiple sclerosis and Sjogren's disease. The
 CC present sequence represents a protein described in relation to green
 CC fluorescent protein (GFP)-fused AChE variant expression construct pGARP
 CC in the examples of the present invention
 XX
 XX SQ Sequence 53 AA;
 XX
 XX The present invention relates to C-terminal peptides of
 CC acetylcholinesterase (AChE) (see AAB5032-B5034). The peptides of the
 CC present invention have cell growth and/or cell differentiation activity.
 CC The peptides may be used in ex vivo or in vivo expansion of
 CC haematopoietic stem cells and neural progenitors, and in the promotion of
 CC megakaryocytic differentiation of hematopoietic stem cells. In addition,
 CC the peptides may be used for promoting expansion of committed neural
 CC progenitors in a developing embryo, in cultured embryonic stem cells, and
 CC embryoid bodies derived from them. The peptides may further be used in
 CC the treatment of thrombocytopenia, post-irradiation conditions, post-
 CC chemotherapy conditions, and conditions following massive blood loss,
 CC inducing synthesis of AChE mRNA, and in promoting formation of hematopoietic
 CC bodies. Antibodies directed against the peptides are useful for
 CC diagnosing stress-induced male infertility. The present sequence is a C-
 CC terminal AChE "readthrough" protein (ARP), which was used in a yeast two-
 CC hybrid system, to screen for ARP binding partners.
 XX
 XX SQ Sequence 53 AA;
 XX
 XX Query Match 100.0%; Score 147; DB 4; Length 53;
 XX Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX Qy 1 GMQGPAGSGWREGSGSPPGTPLFSP 26
 XX Db 28 GMQGPAGSGWREGSGSPPGTPLFSP 53

RESULT 9
 ABG1331
 ID ABG1331 standard; protein; 53 AA.
 XX AC ABG1331;
 AC XX DT 05-NOV-2002 (first entry)
 XX DE GFP-fused AChE variant expression construct, pGARP related protein.
 XX Nervous system; drug assay; acetylcholinesterase; AChE; brain;
 KW isoenzyme variance; AChE blocker; muscarinic receptor; M1; M2;
 KW pyridostigmine; muscarinic receptor blocker; scopolamine;
 KW M1 receptor blocker; pirenzepine; anxiety; post-traumatic stress;
 KW Alzheimer's disease; muscle malfunction; neurodegenerative disorder;
 KW xenobiotic damage; panic; neuromuscular disorder; Parkinson's disease;
 KW Huntington's chorea; muscle fatigue; multiple chemical sensitivity;
 KW autism; multiple sclerosis; Sjogren's disease; GFP; pGARP;
 KW green fluorescent protein.
 OS Unidentified.
 XX WO200240994-A2.
 PN XX PD 23-MAY-2002.
 XX PF 14-NOV-2001; 2001WO-IL001051.
 PR XX 14-NOV-2000; 2000US0-0247970P.
 PA XX

RESULT 10
 AAW48797
 ID AAW48797 standard; protein; 600 AA.
 XX AC AAW48797;
 AC XX DT 07-OCT-1998 (first entry)
 XX DE Human acetylcholinesterase-I4 readthrough splice variant.
 XX KW Human acetylcholinesterase-I4 readthrough splice variant; AChE-I4; CNS;
 KW blood/brain barrier; BBB; I4 peptide; antibiotic; brain tumour; glioma;
 KW chemotherapeutic drug; central nervous system.
 XX OS Homo sapiens.
 KW Key
 KW Region 1..574
 /note= "This region is encoded by exons 1-4 of AChE"
 KW 575..599
 /note= "I4 peptide encoded by intron 4 of AChE; this
 sequence is claimed by the inventors under claim 2 in the
 specification"
 KW 600
 KW Region WO9822132-A2.
 KW XX PD 28-MAY-1998.
 KW XX PF 20-NOV-1997; 97WO-US021196.

PR 20-NOV-1996; 96US-0031194P.
 PR 12-DEC-1996; 96US-0035266P.
 PR 21-JUL-1997; 97US-0053200P.
 XX
 PA (VISS) YISSUM RES & DEV CO.
 PA (KOHN/) KOHN K I.
 PI Soreq H, Friedman A, Seidman S, Kaufer D;
 XX DR; 1998-312172/27.
 XX
 Increasing the permeability of the blood/brain barrier - using e.g.
 PT adrenergine, atropine or acetylcholine esterase I4 splice variant Peptide,
 PT useful For imaging and/or treatment of central nervous system disorders.
 XX
 Claim 1 2; Page 45; 71pp; English.
 XX
 The present sequence represents human acetylcholine esterase-I4 (AChE-I4)
 CC readthrough splice variant. The protein sequence comprises residues
 CC encoded by exons 1-4 of human AChE followed by residues encoded by intron
 CC 4, while the last residue of the protein is encoded by exon 5 of AChE.
 CC The invention provides a pharmaceutical composition, for facilitating
 CC passage of compounds through the blood/brain barrier (BBB), comprising of
 CC AChE-I4, 14 peptide or AChE-I4 analogues together with a pharmaceutically
 CC acceptable carrier. The pharmaceutical composition is claimed to
 CC facilitate a reversible disruption of the BBB allowing transport of
 CC compounds through the BBB. The compounds, e.g. imaging agents,
 CC antibiotics or chemotherapeutic drugs, are claimed to be useful for the
 CC diagnosis and treatment of diseases or disorders of the CNS such as
 CC infections, neurochemical disorders, brain tumours, gliomas, etc
 XX
 Sequence 600 AA;
 'Query Match 100.0%; Score 147; DB 2; Length 600;
 Best Local Similarity 100.0%; Pred. No. 3 2e-09;
 ' Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ' DB
 Qy 1 GMQGPAGSWEBGSGSPGPPGVTPLFSP 26
 Db 575 GMQGPAGSWEBGSGSPGPPGVTPLFSP 600
 RESULT 11
 ID ADC33366 standard; protein; 145 AA.
 XX
 AC ADC33366;
 AC DT 18-DEC-2003 (first entry)
 DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3448.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; anticarneolsonian; nootropics;
 KW neuroprotective; antianemic; antiangiulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy.
 XX
 OS Homo sapiens.
 PN WO2003032271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX DR N-PSDB; Adc32599.
 XX
 New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases; anaemia, platelet
 PR disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PR cancer.
 XX
 PS Example 2; SEQ ID NO 3448; 1185pp; English.
 XX
 The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC3080-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human DNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; and methods of
 CC detecting polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32227) and the polypeptides encoded by the contigs
 CC -ADC33394. The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependant on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases; anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO
 CC at ftp://wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 145 AA;
 'Query Match 42.9%; Score 63; DB 7; Length 145;
 Best Local Similarity 54.2%; Pred. No. 8.7;
 ' Matches 13; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 ' DB
 Qy 3 QGPAGSGWEEGSGSPGPPGVTPLFSP 26
 Db 71 QGDRGTEREGSGSPGPTAGMP 94
 RESULT 12
 ID ABB11475 standard; peptide; 575 AA.
 XX
 AC ABB11475;
 AC DT 11-JAN-2002 (first entry)
 DE Human R31449_3 homologue, SEQ ID NO:1845.
 XX
 KW haematopoiesis regulation; tissue growth; immunomodulator; growth factor;
 KW chemotaxis; chemokines; thrombosis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;

chronic inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischaemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteoprotector; vasotrophic; cardiant; antibacterial;
antifungal; vulnerary; antiulcer.

Homo sapiens.

WO200157188-A2.

XX

PN

XX

PD

XX

09-AUG-2001.

PP

XX

05-FEB-2001: 2001WO-US0031800.

PR

XX

03-FEB-2000: 2000US-00496914.

PR

27-APR-2000: 2000US-00560875.

XX

(HYSS-)

HYSEQ INC.

PA

XX

Tang YT, Liu C, Drmanac RT;

PI

XX

WPI: 2001-457740/49.

DR

N-PSDB; ABA08719.

XX

PT

XX

Human Proteins and DNA encoding sequences useful for preventing, treating

PT

or ameliorating a medical condition in a mammalian subject e.g. arthritis

PT

XX

PS

XX

Claim 20: Page 197; 1961pp; English.

XX

Sequences ABB10981-ABB12330

represent

1350 novel human polypeptides, and

sequences ABA08225-ABA09574

represent

nucleic acids

encoding

them. The

invention

also

relates

to

vectors

and

recombinant

host

cells

comprising

a

nucleotide

of

the

novel

polypeptides,

methods

of

detecting

the

nucleotides

or

polypeptides

in

a

sample,

and

methods

of

identifying

compounds

which

bind

to

polypeptides

of

the

invention.

Although

novel,

many

of

the

polypeptides

of

the

invention

have

homology

to

known

proteins,

thereby

giving

an

insight

into

their

probable

biological

activities,

and

hence

potential

therapeutic

applications.

The

polypeptides

of

the

invention

may

have

various

activities

including

cytokine,

cell

proliferation

or

metastasis.

Depending

on

their

biological

activities,

polypeptides

and

nucleotides

of

the

invention

are

useful

for

preventing,

treating

or

ameliorating

medical

conditions

including

cancers,

hematopoietic

disorders

(e.g.,

myeloid

or

lymphoid

cell

disorders),

chronic

inflammatory

conditions

(e.g.,

asthma

or

arthritis),

proliferative

retinopathy,

atherosclerosis,

coronary

heart

disease,

arterial

ischaemia,

bone

disorders

(e.g.,

osteoporosis),

abnormal

vascular

growth.

Polypeptides

involved

with

tissue

regeneration

and

repair

(or

nucleic

acids

encoding

them)

may

be

used

to

promote

wound

healing

(e.g.,

of

burns,

incisions

and

ulcers),

while

those

with

immunomodulatory

activities,

receptor

or

ligand

activities;

or

involved

in

oncogenesis,

cancer

cell

proliferation

or

metastasis.

Depending

on

their

biological

activities,

polypeptides

and

nucleotides

of

the

invention

are

useful

for

preventing,

treating

or

ameliorating

medical

conditions

including

cancers,

hematopoietic

disorders

(e.g.,

myeloid

or

lymphoid

cell

disorders),

chronic

inflammatory

conditions

(e.g.,

asthma

or

arthritis),

proliferative

retinopathy,

atherosclerosis,

coronary

heart

disease,

arterial

ischaemia,

bone

disorders

(e.g.,

osteoporosis),

abnormal

vascular

growth.

Polypeptides

involved

with

tissue

regeneration

and

repair

(or

nucleic

acids

encoding

them)

may

be

used

to

promote

wound

healing

(e.g.,

of

burns,

incisions

and

ulcers),

while

those

with

immunomodulatory

activities,

receptor

or

ligand

activities;

or

involved

in

oncogenesis,

cancer

cell

proliferation

or

metastasis.

Depending

on

their

biological

activities,

polypeptides

and

nucleotides

of

the

invention

are

useful

for

preventing,

treating

or

CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional
 CC components. Oligonucleotides AAC93470 - AAC93478 and Peptide AAB51826 are
 CC used in the isolation and characterisation of the proteins and
 CC polynucleotides of the invention

XX SQ Sequence 98 AA;

Query Match 41.5%; Score 61; DB 3; Length 98;
 Best Local Similarity 59.1%; Pred. No. 10;
 Matches 13; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy 5 PAGSGMVEEGSGSPPGVTPFLSP 26
 Db 64 PRGSGNERAPGCP--VTPLTLP 83

RESULT 14

AAR37744 standard; protein; 54 AA.
 ID AAR37744

XX AAR37744;

XX 25-MAR-2003 (revised)

DT 07-SEP-1993 (first entry)

XX Collagen-like polymer DCP - (DB) 3.

DE Recombinant; collagen-like polymer; CLP; tripeptide; helix; membrane;
 KW fibre; film; coating; triad sequence; collagen; mammalian; moulding;
 KW hydrogel; interchain linkage; colloid suspension; DCP; antibody.
 XX Synthetic.

XX OS Synthetic.

XX WO9310154-A1.

XX PD 27-MAY-1993.

XX PP 04-NOV-1992;

XX PR 12-NOV-1991;

XX PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX PI Cappello J, Ferrari FA;

XX PN WO9310154-A1.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX The invention concerns collagen-like polymers having repetitive triads
 CC with reduced proline content, and where Glycine is the initial amino acid
 CC and the subsequent amino acids are varied. The choice of triads utilised
 CC in a recombinant collagen-like polymer are chosen in order to affect
 CC properties such as helix stability, hydration, solubility, gel point,
 CC biodegradation and immunogenicity. Triads of particular interest include
 CC GAP, GPA, GPP, GAS, GPG, GPS, GLO, GPR, GPK, GAK, GAR,
 CC GDR, GEP, GDA, GAH and GEA. The collagen-like polymers may impart new
 CC characteristics, finding wide use in photographic, medical, structural
 CC and fibre applications, and are capable of being produced in unicellular
 CC microorganisms at high mol. wts. and in high efficiency. The present
 CC sequence, encoded by clone pTO224 (see AAT16769), was identified to
 CC contain the sequenced gene 4 or 5 monomer sequence (DB) 3. The sequenced
 CC collagen-like polymers are used as immunogens for the prepn. of
 CC antibodies. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-
 CC MAR-2003 to correct PA field.)

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX The sequences given in AAR37744 and AAR37747 represent examples of
 CC recombinantly produced DCP collagen-like polymers (CLPs) which consist of
 CC repeated tripeptide sequences selected from a wide range of GXY
 CC sequences, where X and Y can be any amino acid. The DNA encoding these
 CC sequences can be cloned into plasmids and used to transform E. coli to
 CC produce the DCP proteins. DCP peptides comprise repeated units of; B =
 CC GSRGDPEPP, C = GAGPAGPK and/or D = GAQGAPGP. These polymers have
 CC molecular weights of >30 kD and are able to form helices due to
 CC interchain linkages. These polymers pref. contain a proportion of
 CC tripeptide triad sequences found in natural collagens, pref. mammalian
 CC collagens. The CLPs impart unique characteristics to materials such as
 CC fibres, membranes, films, coatings, hydrogels, colloid suspensions and
 CC moulded articles. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 54 AA;

Query Match 40.8%; Score 60; DB 2; Length 54;
 Best Local Similarity 57.9%; Pred. No. 7.3;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Query Match 40.8%; Score 60; DB 2; Length 54;
 Best Local Similarity 57.9%; Pred. No. 7.3;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Search completed: October 29, 2005, 01:03:13

Qy 1 GMQGPAGSGMVEEGSGSPPG 19
 Db 1 GAQGPAGSGMVEEGSGSPPG 19

RESULT 15

AAR3255 standard; protein; 54 AA.
 ID AAR3255

XX AAR3255;

XX AC AAR3255;

XX DT 25-MAR-2003 (revised)

DT 24-FEB-1997 (first entry)

XX DE Collagen-like polymer sequence D4/DS unit (DB) 3.

XX KW collagen; repetitive triad motif; recombinant production; photographic;

XX KW medical; structural; fibre.

XX OS Synthetic.

XX PN US5496712-A.

XX PD 05-MAR-1996.

XX PP 05-NOV-1992;

XX PR 06-NOV-1990;

PR 12-NOV-1991;

XX PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX PI Cappello J, Ferrari FA;

XX PN WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g.

XX PT photographic and medical fibres.

XX PS Example 3; Col 23-24; 43pp; English.

XX Sequence 54 AA;

XX Disclosure; Page 52; 82pp; English.

XX PS Disclosure; Page 52; 82pp; English.

XX DR WPI:1996-150728/15.

XX DR N-PSDB; AAT16769.

XX PT Collagen-like polymers comprising repetitive triads - produced in

XX PT unicellular organisms with improved characteristics, useful in, e.g

A;Title: The covalent structure of calf skin type III collagen. III. The amino acid sequence of the carboxy-terminal nonhelical telopeptide of type III collagen. A;Reference number: A38002; MUID:80026028; PMID:488908

A;Accession: A38002

A;Molecule type: protein

A;Residues: 423-571 <BEN>

R;Lang, H.; Gianville, R.W.; Fietzek, P.P.; Kuehn, K.

Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979

A;Title: The covalent structure of calf skin type III collagen. IV. The amino acid sequence of the carboxy-terminal nonhelical telopeptide of type III collagen. A;Reference number: A38003; MUID:80026029; PMID:488909

A;Accession: A38003

A;Molecule type: protein

A;Residues: 577-808 <LAN>

R;Davies, H.; Fietzek, P.P.; Kuehn, K.

Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979

A;Title: The covalent structure of calf skin type III collagen. V. The amino acid sequence of the carboxy-terminal nonhelical telopeptide of type III collagen. A;Reference number: A38004; MUID:80026030; PMID:488910

A;Accession: A38004

A;Molecule type: protein

A;Residues: 809-947 <DEV>

R;Altmann, H.; Fietzek, P.P.; Kuehn, K.

Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979

A;Title: The covalent structure of calf skin type III collagen. VI. The amino acid sequence of the carboxy-terminal nonhelical telopeptide of type III collagen. A;Reference number: A38005; MUID:80026031; PMID:488911

A;Accession: A38005

A;Molecule type: protein

A;Residues: 948-1049 <ALI>

A;Experimental source: skin

R;Hankel, W.

Biochem. J. 318, 497-503, 1996

A;Title: Cross-link analysis of the C-telopeptide domain from type III collagen.

A;Reference number: S71946; MUID:95404837; PMID:8809038

A;Accession: S71946

A;Molecule type: protein

A;Residues: 87-106-1017-1029-1037-1049 <HEN>

C;Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are linked to each other via a disulfide bond. The repeating units are linked to each other via a disulfide bond.

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxy-terminal homology

C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline; hydroxyproline; collagen alpha 1(III) chain #status experimental

F;1-1049/Region: amino-terminal nonhelical telopeptide

F;1-1040/Region: helical

F;567-589/Region: cell attachment (R-G-D) motif

F;742-754/Region: cell attachment (R-G-D) motif

F;857-877/Region: cell attachment (R-G-D) motif

F;878-880/Region: cell attachment (R-G-D) motif

F;955-937/Region: cell attachment (R-G-D) motif

F;1041-1049/Region: carboxyl-terminal nonhelical telopeptide

F;107,950/Modified site: 5-hydroxylysine (Lys) #status predicted

F;107/Binding site: carbohydrate (Lys) (covalent) #status experimental

F;1040,1041/Disulfide bonds: interchain #status predicted

Query Match Score 40 %; Best Local Similarity 57.1%; Matches 12; Conservative 0; Mismatches 0; Gaps 0; Result 5

Db 447 GPQGPGKNGKGTGPQGPGPPT 467

RESULT 3

A;Title: The covalent structure of calf skin type III collagen. III. The amino acid sequence of the carboxy-terminal nonhelical telopeptide of type III collagen. A;Reference number: A41573; MUID:1744105

A;Accession: A41573

A;Molecule type: protein

A;Residues: 30-46 'R, 48-72 <LIV>

A;Cross-references: GB:M76539; NID:9177976; PID:9553165

A;Status: preliminary

A;Gene: GDB:ACHE

A;Cross-references: GDB:118746; OMIM:100740

A;Keywords: alternative splicing; carboxylic ester hydrolase

Query Match Score 39.8%; Best Local Similarity 50.0%; Matches 14; Conservative 2; Mismatches 9; Indels 3; Gaps 1; Result 4

Db 1 GMQGPAGSGCWEEGSGS---PPGVTPLPS 25

Db 4 GMQGPAGSGAQRGVGARQNPSPPLPS 31

RESULT 4

JC2217

A;Title: Major surface glycoprotein 5 - Pneumocystis carinii

A;Accession: JC2217

A;Cross-references: R;Kitada, K.; Wada, M.; Nakamura, Y.

C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004

C;Accession: JC2217

A;Title: Multi-gene family of major surface glycoproteins of Pneumocystis carinii: full-length cDNA clones and sequence analysis. A;Reference number: JC2217; MUID:7584029

A;Accession: JC2217

A;Molecule type: mRNA

A;Residues: 1-1076 <KIT>

A;Cross-references: UNIPROT:Q01830; DDBJ:D21827; NID:9425784; PID:d1005

C;Superfamily: Pneumocystis carinii major surface glycoprotein MSG100

C;Keywords: Glycoprotein

F;245,471,574,804,837/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match Score 39.8%; Best Local Similarity 65.0%; Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1; Result 5

Db 3 QGPAGSGCWEEGSGSPPGVT 22

Db 816 QAPAGS--SGGSSPPAVPP 832

RESULT 5

A;Title: Hypothetical protein At2g35130 [imported] - Arabidopsis thaliana

A;Accession: A84765

A;Cross-references: C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: A84765

R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umam, L.; Salzberg, S.L.; Fraser, C.M.; Venler, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Accession: A84420

A;Cross-references: UNIPROT:O82178; GB:AE002093; NID:93668091; PID:20083487; MUID:10617197

A;Molecule type: DNA

A;Residues: 1-591 <STO>

A;Cross-references: UNIPROT:O82178; GB:AE002093; NID:93668091; PID:20083487; MUID:10617197

A;Genetics: At2g35130

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Map position: 2

Query Match 39.5%; Score 58; DB 2; Length 591;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 7 GSGWEGSGSPPGVTPFLFSP 26

Db 68 GRGWKYGSVPFGVDFGIFVLSLSP 87

RESULT 6

collagen alpha 1(XI) chain - chicken (fragment)

C;Species: Gallus gallus (chicken)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: 28791
R;Nah, H.D.; Barenbaum, M.; Upholt, W.B.
J. Biol. Chem. 267, 22581-22586, 1992
A;Title: The chicken alpha1 (XI) collagen gene is widely expressed in embryonic tissues.
A;Reference number: S28791; MUID:93054557; PMID:14296607
A;Accession: 28791
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-888 <NAH>
A;Cross-references: UNIPROT:Q90796; EMBL:M88593; NID:9211619; PIDN:AAA48707.1; PID:g2116
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology <FCC>
F:655-887/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 39.5%; Score 58; DB 2; Length 888;
Best Local Similarity 53.8%; Pred. No. 22;
Matches 14; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 26

Db 118 GAQGPAGLKCEGPGQPGP-PVGSP 141

RESULT 7

collagen alpha 1(III) chain precursor - mouse

C;Species: Mus musculus (house mouse)
C;Date: 10-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C;Accession: S59856; S6120; S61373
R;Toman, P.D.; de Crombrugge, B.
Gene 147, 161-168, 1994
A;Title: The mouse type-III procollagen encoding gene: genomic cloning and complete DNA
A;Accession: S59856; MUID:95011609; PMID:7926795
A;Molecule type: DNA
A;Residues: 1-1464 <TOM>
A;Cross-references: UNIPROT:P08121; EMBL:X52046
R;Toman, P.
A;Reference number: S62120
A;Accession: S62120
A;Molecule type: DNA
A;Cross-references: EMBL:X52046; NID:95753221; PIDN:CAA36279.1; PID:95753222
R;Metsaeranta, M.; Toman, D.; de Crombrugge, B.; Vuorio, E.
Biochim. Biophys. Acta 1089, 241-243, 1991
A;Accession: S16373
A;Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
A;Reference number: S16376; MUID:91274355; PMID:2054384
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 142-1464 <NET>
A;Cross-references: EMBL:X57983; NID:950476; PIDN:CAA41048.1; PID:950477

C;Genetics:
C;Introns: 29/1; 95/3; 112/3; 150/3; 175/3; 193/3; 211/3; 229/3; 247/3; 265/3; 283/3; 295/3; 672/3; 706/3; 742/3; 760/3; 778/3; 795/3; 814/3; 850/3; 868/3; 940/3; 976/3
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C;Keywords: coiled coil; extracellular matrix

F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-154/Domain: propeptide #status predicted <PRO>
F:12-92/Domain: von Willebrand factor type C repeat homology <WFC>
F:155-1464/Domain: collagen alpha 1(III) chain #status predicted <MAT>
F:1236-1464/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 39.5%; Score 58; DB 2; Length 1464;
Best Local Similarity 52.4%; Pred. No. 37;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 21

Db 602 GLPSPAGKGETGPQGPepAPT 622

RESULT 8

hypothetical protein C09G5.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19143
R;Palmer, S.
A;Submitted to the EMBL Data Library, November 1994
A;Accession: T19143
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-317 <WIL>
A;Cross-references: UNIPROT:Q09456; EMBL:246791; PIDN:CAA867558.1; GSPDB:GN00020; CBSP:CO

Query Match 39.5%; Score 57; DB 2; Length 317;
Best Local Similarity 57.9%; Pred. No. 10;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 229

Db 211 GPQGPAGKGETGPQGPepAPT 229

RESULT 9

collagen alpha 2 (VIII) chain - human (fragment)

C;Species: Homo sapiens (man)
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: A57131
R;Muragaki, Y.; Jacewko, O.; Apre, S.; Mattei, M.G.; Ninomiya, Y.; Olsen, B.R.
J. Biol. Chem. 266, 7721-7727, 1991
A;Title: The alpha2 (VIII) collagen gene: A novel member of the short chain collagen family
A;Reference number: A57131; MUID:91210292; PMID:2013595
A;Accession: A57131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-635 <MDR>
A;Cross-references: UNIPROT:P25067; GB:M60832; NID:9177178; PID:9177179
C;Genetics:
A;Gene: GDB:COL8A2
A;Cross-references: GDB:127812; OMIM:120252
A;Map position: 1234.3-1p32.3
C;Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
F:1-111/Domain: amino-terminal (fragment) #status predicted <NC2>
F:12-468/Region: interrupted helical
F:469-635/Domain: carboxyl-terminal nonhelical #status predicted <NC7>
F:508-634/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match 38.8%; Score 57; DB 2; Length 635;
Best Local Similarity 55.0%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

A;Reference number: 154355; MUID:95187161; PMID:7881420	Db	766	GPNGPPGAGSGRGDDGPPGWT 786
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: mRNA			
A;Residues: 663-675, 'P', 677, 'P', 679-742, 'A', '744-746 <FOR>			
A;Cross-references: GB:L47668; NID:91009095; PIDN:ABA59377.1; PID:91009096			
A;R;Niyibizi, C.; Bonadio, J.; Byers, P.H.; Eyrz, D.R.			
A;J. Biol. Chem. 267, 23108-23112, 1992			
A;Title: Incorporation of type I collagen molecules that contain a mutant alpha 2(I) chain into the extracellular matrix of a patient with osteogenesis imperfecta			
A;Reference number: 155369; MUID:93054637; PMID:13885413			
A;Status: translated from GB/EMBL/DBJ			
A;Molecule type: mRNA			
A;Residues: 665-666, 'D', 668-670 <NFT>			
A;Cross-references: GB:L00613; NID:91008088; PIDN:ABA59384.1; PID:91008089			
A;Note: mutant sequence from a patient with osteogenesis imperfecta			
A;R;Batemann, J.F.; Rannung, M.; Chan, D.; Cole, W.G.			
A;J. Biol. Chem. 276, 765-770, 1991			
A;Title: Characterization of a type I collagen alpha 2(I) glycine-586 to valine substitution in a patient with osteogenesis imperfecta			
A;Reference number: A56799; MUID:91291136; PMID:2064612			
A;Accession: A56799			
A;Molecule type: mRNA			
A;Residues: 672-675, 'V', 677, 'P', 679-681 <BAT>			
A;Cross-references: GB:S30878; NID:91679911; PIDN:ABA19314.1; PID:9232761			
A;Note: sequence extracted from NCBI backbone (NCBIN:39878, NCBI:39886)			
A;Reference number: S10768; MUID:90304220; PMID:2364107			
A;Molecule type: mRNA			
A;Residues: 960-1021, 'L', 1023-1188, 'D', 1190-1197, 'S', 1199-1356 <MAE>			
A;Cross-references: EMBL:X55552; NID:930101; PIDN:CAA39142.1; PID:930102			
A;Experimental source: fibroblast cell culture			
A;R;Myers, J.C.; Chu, M.L.; Faro, S.H.; Clark, W.J.; Prockop, D.J.; Ramirez, F.			
A;Proc. Natl. Acad. Sci. U.S.A. 78, 3516-3520, 1981			
A;Title: Cloning a cDNA for the pro-alpha2 chain of human type I collagen.			
A;Reference number: A18855; MUID:81273090; PMID:6267597			
A;Accession: A18855			
A;Molecule type: mRNA			
A;Residues: 961-979, 'V', 981-1018, 'Q', 1020 <MYE>			
A;A;Note: 1019-Leu was also found			
A;R;Wenstrup, R.J.; Cohn, D.H.; Cohen, T.; Byers, P.H.			
A;J. Biol. Chem. 263, 7734-7740, 1988			
A;Title: Arginine for glycine substitution in the triple-helical domain of the products of the human pro-alpha2 chain of collagen.			
A;Reference number: 155283; MUID:882327975; PMID:2837363			
A;Status: translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1050-1101, 'R', 1103-1107 <WEN1>			
A;Cross-references: GB:M22816; NID:9179602; PIDN:AAA51844.1; PID:9179603			
A;Accession: I70059			
A;Molecule type: DNA			
A;Residues: 1050-1101, 'R', 1103-1107 <WEN2>			
A;Cross-references: GB:M2817; NID:9179606; PIDN:AAA51846.1; PID:9179607			
A;A;Note: mutant sequence from a patient with osteogenesis imperfecta type IV			
A;R;Myers, J.C.; Dickson, L.A.; de Wet, W.J.; Bernard, M.P.; Chu, M.L.; di Liberto, M.; Pepeu, G.			
A;J. Biol. Chem. 258, 10128-10135, 1983			
A;Title: Analysis of the 3' end of the human pro-alpha2(I) collagen gene. Utilization of the 3' end of the human pro-alpha2(I) collagen gene.			
A;Reference number: S09175; MUID:83290853; PMID:6309769			
A;Accession: S09175			
A;Molecule type: DNA			
RESULT 13			
T18594			
hypothetical protein AC3.6 - Caenorhabditis elegans			
C;Species: Caenorhabditis elegans			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1994			
C;Accession: T18594			
C;R;McMurray, A.			
A;Reference number: 218995			
A;Submitted to the EMBL Data Library, April 1996			
Query Match 38.8%; Score 57; DB 1; Length 1366;			
Best Local Similarity 47.6%; Pred. No. 46;			
Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;			
1 GMQGPAGSGWEGSGSPPGVT 21			
150 GLVGPAGPAGDQGRHGPBPGBT 170			
RESULT 14			
T18594			
hypothetical protein AC3.6 - Caenorhabditis elegans			
C;Species: Caenorhabditis elegans			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1994			
C;Accession: T18594			
C;R;McMurray, A.			
A;Reference number: 218995			
A;Submitted to the EMBL Data Library, April 1996			
Query Match 38.8%; Score 57; DB 1; Length 1366;			
Best Local Similarity 47.6%; Pred. No. 46;			
Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;			
1 GMQGPAGSGWEGSGSPPGVT 21			
150 GLVGPAGPAGDQGRHGPBPGBT 170			

A;Accession: T18594
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-325
 A;Cross-references: UNIPROT:Q17402; EMBL:Z71177; PIDN:CAA94869.1; GSPDB:GN00023; CESP:AC3
 A;Experimental source: clone AC3
 C;Genetics:
 A;Gene: CESP:AC3.6
 A;Map Position: 5
 A;Introns: 47/3; 101/1; 270/3
 Query Match Score 56; DB 2; Length 325;
 Best Local Similarity 57.9%; Pred. No. 14;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 1 GMQGPAGSGWEEGGSGSPPG 19
 Dbs 218 GSPGPAGSPGRVGRGPPG 236

RESULT 14

B34493
 collagen alpha 1(IX) chain long form precursor - chicken
 C;Species: Gallus gallus (chicken)

B34493; #sequence 28-Oct-1994 #text change 09-Jul-2004
 C;Accession: B34493; S21861; B28360; S22243; A20982; A28754; A30973; S22240
 R;Nishimura, I.; Muragaki, Y.; Olsen, B.R.
 J; Biol. Chem. 264, 20033-20041, 1989
 A;Title: Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two
 A;Reference number: A31493; MUID:0062114; PMID:2584206
 A;Molecule type: mRNA
 A;Residues: 1-416 <NIS>
 A;Cross-references: UNIPROT:PI2106; UNIPROT:Q90779; GB:M28659; NID:9211623; PIDN:AAA4874
 R;Olsen, B.R.
 A;Reference number: B34493
 A;Accession: S21861
 A;Submitted to the ENBL Data Library, December 1987
 A;Reference number: S21861
 A;Accession: S21861
 A;Molecule type: mRNA
 A;Residues: 1-24; 'I', 26-50, 'Q', 52-266 <VAS>
 A;Cross-references: EMBL:J03539
 A;Molecule type: protein
 A;Residues: 31-38 'I'; 49-50, 'Q', 52-54; 97-114; 152-259-266 <VAA>
 R;Nishimura, I.; Konon, H.; van der Reest, M.; Niniomiya, Y.; Olsen, B.R.
 J; Biol. Chem. 263, 2324-2329, 1988
 A;Title: Cartilage type IX collagen-proteoglycan contains a large amino-terminal globular
 A;Reference number: A28360; MUID:88115376; PMID:3339014
 A;Accession: A28360
 A;Molecule type: mRNA
 A;Residues: 'I', 26-50, 'Q', 52-266
 A;Cross-references: B28360
 A;Molecule type: protein
 A;Residues: 31-38 'I'; 49-50, 'Q', 52-54; 97-114; 152-259-266 <VAA>
 R;Nishimura, I.; Olsen, B.R.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McCollum, R.; Lozano, G.; Niniomiya, Y.; Thompson, H.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 3014-3018, 1984
 A;Title: Synthesis and characterization of collagen genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Press
 A;Reference number: S22243
 A;Accession: S22243
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-24; 'I', 26-50, 'Q', 52-266 <NII>
 A;Cross-references: EMBL:K01702; PIDN:AAA48675.1; PMID:9211499
 R;Lozano, G.; Niniomiya, Y.; Thompson, H.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4050-4054, 1985
 A;Title: A distinct class of vertebrate collagen genes encodes chicken type IX collagen
 A;Accession: A20982
 A;Molecule type: mRNA
 A;Residues: 183-920 <NII>
 A;Cross-references: EMBL:K01702; PIDN:AAA48675.1; PMID:9211499

Query Match 38.1%; Score 56; DB 1; Length 1373;
Best Local Similarity 48.0%; Pred. No. 61;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
Qy 1 GMQGPAGSGWEGSSSPGGTPLFS 25
Db 60 GPPGPAGSGPAPGSPAPGUTGNPA 84

Search completed: October 29, 2005, 01:03:31
Job time : 17 secs

DT	DE	Collagen al (I).	26.	Last annotation update)
GN	Name=COL1A1;			
OS	Oncothrychus mykiss	(Rainbow trout)	(<i>Salmo gairdneri</i>).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.			
OX	NCBI_TaxID=8022;			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=21257802; PubMed=11358497;			
RX	Saito M., Takenouchi Y., Kunisaki N., Kimura S.;			
RA	Complete primary structure of rainbow trout type I collagen			
RT	consisting of all(I) a2(I) a3(I) heterotrimers.";			
ER	Eur. J. Biochem. 268:2817-2827(2001).			
RL	EMBL; AB052355; BAB55661.1.			
DR	GO: GO:000581; C:collagen; IEA.			
DR	GO: GO:0005237; C:cytoplasm; IEA.			
DR	GO: GO:0005201; F:extracellular matrix structural constituent; IEA.			
DR	PFAM; PF01410; COLFI; 1.			
DR	PFam; PF01391; Collagen; 18.			
DR	PFam; PF00093; VWC; 1.			
DR	ProDom; PDD00007; CIG_helix; 4.			
DR	ProDom; PDD02078; Fib_collagen_C; 1.			
DR	SMART; SM00038; COLIFI; 1.			
DR	SMART; SM000214; VWC; 1.			
DR	PROSITE; PS01208; WFPC_1; 1.			
DR	PROSITE; PS50184; WFPC_2; 1.			
KW	Collagen.			
SEQUENCE	1449 AA; 137116 MW; 62EEF8A7BF652B8 CRC64;			
Query Match	45..6%; Score 67; DB 2; Length 1449;			
Best Local Similarity	46..2%; Pred. No. 27;			
Matches 12;	Conservative 4; Mismatches 10; Indels 0; Gaps 0;			
Qy	1 GMQGPAGSGWEEGSSGSPICVTPFLFSP 26			
Db	608 GYVAPGSSERGEQAGGPPGFQGLSGP 633			
RESULT 3				
Q6BP55	Q6BP25; PRELIMINARY;	234 AA.		
AC	Q6BP25; 25-OCT-2004 (TREMBLrel. 28, Created)			
DT	DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)			
DT	DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE	Similarity			
RA	Goffard N., Frangeul L., Aigle M., Marck C., Neuveglise C., Talia B., Durrans P., Casaregola S., Lafontaine I., de Montigny J., March C., Babbou A., Barbe V., Boisrame A., Blanchard J.M., Beyne E., Bleykasten C., Boisrame A., Boyer J., Cattecoile L., Confaoliolier F., Desponts L., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye F., Hennequin C., Jauziaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicand J.M., Nitolski M., Oztas S., Ozier-Kalogeropoulos O., Pellegrin S., Potier S., Richard G.F., Straub M.L., Suleau A., Swennenc D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Weissenbach J., Wincker P., Souciet J.L.;			
RN	SEQUENCE FROM N.A.			
RP	STRAN=CBS767;			
RG	Genelavers;			
RA	Goffard N., Frangeul L., Aigle M., Marck C., Neuveglise C., Talia B., Durrans P., Casaregola S., Lafontaine I., de Montigny J., March C., Babbou A., Barbe V., Boisrame A., Blanchard J.M., Beyne E., Bleykasten C., Boisrame A., Boyer J., Cattecoile L., Confaoliolier F., Desponts L., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye F., Hennequin C., Jauziaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicand J.M., Nitolski M., Oztas S., Ozier-Kalogeropoulos O., Pellegrin S., Potier S., Richard G.F., Straub M.L., Suleau A., Swennenc D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Weissenbach J., Wincker P., Souciet J.L.;			
RA	Genome evolution in yeasts.";			
RA	Nature 430:35-44 (2004).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CBS767;			
RA	Genoscope;			
RA	Submitted (MMI-2004) to the EMBL/GenBank/DDJB databases.			
RL	EMBL; CR88137; CAG88303.1; -			
DR	GO: GO:0005737; C:cytoplasm; IEA.			
DR	GO: GO:0006817; P:phosphate transport; IEA.			
DR	InterPro; IPR008161; CIG_helix.			
DR	InterPro; IPR008160; Collagen.			
DR	PFam; PF01391; Collagen. 3.			
DR	ProDom; PDD00007; CIG_helix; 1.			
KW	Collagen.			
SEQUENCE	234 AA; 22636 MW; 3C0B851B2D0095C9 CRC64;			
QY	1 GMQGPAGSGWEEGSSGPPGV 20	42.2%; Score 62; DB 2; Length 234;		
Db	126 GDQGPXGAGGERGIGGPGQGV 145	60.0%; Pred. No. 16; Mismatches 7; Indels 0; Gaps 0;		
RESULT 4				
Q6KAQ4	Q6KAQ4; PRELIMINARY;	640 AA.		
ID	Q6KAQ4; 05-JUL-2004 (TREMBLrel. 27, Created)			
AC	Q6KAQ4; 05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT	DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DB	DB MFL00201 protein (Fragment). GN Name=mFL00201;			
OS	OS Mus musculus (Mouse).			
OC	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Mus. NCBI_TaxID=10090;			
RN	RN [1] _TAXID=10090;			
RP	SEQUENCE FROM N.A.			
RC	RC TISSUE=Embryonic tail;			
RA	RA Okazaki N., Kitakubo R., Ohara R., Inamoto S., Koseki H., Hiraoka S., Koga H.; Nagase T., Ohara O., Koga H.;			
RA	RA Saga Y., Kitamura H., Nakagawa T., Nakagawa T., Nagase T., Ohara O., Koga H.;			
RT	RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes: The Complete Nucleotide Sequences of 110 Mouse FLJ Homologous cDNAs Sampled from Screening of Terminal Sequences of cDNA Clones Randomly DNA Res. 11:167-180 (2004)."			
RT	RT EMBL; AK131153; BAD1403.1; -			
DR	DR GO: GO:0005737; C:cytoplasm; IEA.			
DR	DR GO: GO:0006817; P:phosphate transport; IEA.			
DR	DR InterPro; IPR01073; CIG_helix.			
DR	DR InterPro; IPR008160; Collagen.			
DR	DR InterPro; IPR00983; TNF-like.			
DR	DR PF00386; CIG_1.			
DR	DR PFam; PF01391; Collagen; 8.			
DR	DR PRINTS; PR00007; COMPLENNTC1Q.			
DR	DR SMART; SM00110; C1Q; 1.			
DR	DR PROSITE; PS01113; C1Q; 1.			
KW	KW Collagen.			
SEQUENCE	640 AA; 161034 MW; 75CC9DEBAE5AC4B5 CRC64;			
QY	1 GMQGPAGSGWEEGSSGPPGV 20	42.2%; Score 62; DB 2; Length 640;		
Db	307 GRRGPDSKGEVGPGPQPGV 326	60.0%; Pred. No. 45; Mismatches 12; Conservative 1; Indels 7; Gaps 0;		

RESULT 5

RL	Bur. J. Biochem.	268:12817-2827(2001).
DR	EMBL; AB02836; BABB5662; 1;	-
DR	GO; GO:0005581; C:collagen; IEA.	
DR	GO; GO:005737; C:cytoplasm; IEA.	
DR	GO; GO:005201; F:extracellular matrix structural constituent; IEA.	
DR	GO; GO:0006817; P:phosphate transport; IEA.	
DR	Pfam; PF04140; COLPF; 1.	
DR	Pfam; PF01391; Collagen; 18.	
DR	Prodom; PRO02078; CIG_helix; 2.	
DR	Prodom; PRO00007; CIG_helix; 1.	
DR	SMART; SM00038; COLPF; 1.	
DR	SMART; SM0214; VWC; 1.	
DR	PROSITE; PS01208; VWC; 1.	
DR	PROSITE; PS50184; VWF_C; 1.	
KW	Collagen.	
SEQUENCE	1458 AA;	137757 MN; AB1F9F3410A98650 CRC64;
Query Match	42.2%; Score 62; DB 2; Length 1458;	
Best Local Similarity	63.2%; Pred. No. 1.1e+02;	
Matches	12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 CMQGPAGSGWBGSGSPPG 19	
Db	611 GPGQGPAGSGQKGEKGSPAG 629	
RESULT 8	FCN1_MOUSE	STANDARD;
ID	070165;	PRT;
AC	AC_070165;	334 AA.
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	15-JUL-2004 (Rel. 44, Last annotation update)	
DE	Ficolin-1 precursor (Collagen/fibrinogen domain-containing protein 1)	
DE	(Ficolin-1) (Ficolin A) (M-Ficolin).	
GN	Name=FCN1; Synonyms=FCNA;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
RA	Pujimori Y., Harumori S., Fukumoto Y., Miura Y., Yagasaki K., Tachikawa H., Fujimoto D.; Precursor (Collagen/fibrinogen domain-containing protein 1)	
RA	"Molecular cloning and characterization of mouse ficolin-A."	
RA	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=BALB/c; TISSUE=Liver;	
RX	MEDLINE=90205801; PubMed=10.1006/bbrc.1998.8344;	
RA	Rahman K., Marusina K., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lognelli N.A., Peters G.J., Abramson R.D., Mullahay S.J., Rosak S.A., McEvlan P.J., McErlean K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehre J., Helton E., Kettman M., Madden A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G., RA	
RA	Rodriguez A.C., Grinowicz J., Krzywinski M.I., Skalska J., Smailus D.E., Butterfield Y.S.N., Schein J.E., Jones S.J.M., Marra M.A., Schnarch A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences," Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RT	FUNCTION: Involved in serum exerting lectin activity. Binds GlnNAc (By similarity).	
RL	SEQUENCE OF 1-242. MEDLINE=80026026; PubMed=488907; Dewes H., Fietzek P.P., Kuhn K.,	
RP	SEQUENCE OF 1-222. MEDLINE=80026027; PubMed=488907; Dewes H., Fietzek P.P., Kuhn K.,	
RX	"The covalent structure of calf skin type III collagen. I. The amino acid sequence of the amino terminal region of the alpha 1(III) chain (positions 1-222)." Hoppe-Seyler's Z. Physiol. Chem. 360:809-820 (1979).	
RN	SEQUENCE OF 1-242. MEDLINE=80026026; PubMed=488907; Dewes H., Fietzek P.P., Kuhn K.,	
RX	SEQUENCE OF 1-242. MEDLINE=80026027; PubMed=488907; Dewes H., Fietzek P.P., Kuhn K.,	

RT "The covalent structure of calf skin type III collagen. II. The amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB1, 8-10, 2 (positions 223-402)." ;
 RT Hoppe-Seyler's Z. Physiol. Chem. 360:821-832(1979).
 [3]
 SEQUENCE OF 423-571.
 RP MEDLINE=80026028; PubMed=488908;
 RX Bentz H., Fietzek P.P., Kuhn K.;
 RA "The covalent structure of calf skin type III collagen. III. The amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB4 (positions 403-551)." ;
 RT Hoppe-Seyler's Z. Physiol. Chem. 360:833-840(1979).
 [4]
 RN
 RP SEQUENCE OF 572-808.
 RX MEDLINE=88026029; PubMed=88909;
 RA Lang H., Gianville R. W., Fietzek P. P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. IV. The amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB5 (positions 552-788)." ;
 RT Hoppe-Seyler's Z. Physiol. Chem. 360:841-850(1979).
 [5]
 RN
 RP SEQUENCE OF 809-947.
 RX MEDLINE=80026030; PubMed=488910;
 RA Dewes H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. V. The amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB9A (position 789-927)." ;
 RT Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).
 [6]
 RN
 RP SEQUENCE OF 948-1049.
 RX MEDLINE=80026031; PubMed=488911;
 RA Allmann H., Fietzek P.P., Gianville R.W., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. VI. The amino acid sequence of the carboxy-terminal cyanogen bromide peptide alpha 1(III)CB9B (positions 928-1028)." ;
 RT Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).
 CC !- FRUNCT: Collagen type III occurs in most soft connective tissues along with type I collagen.
 CC !- SIMILARITY: Trimers of identical alpha 1(III) chains. The chains are linked to each other by interchain disulfide bonds. Trimers are also cross-linked via hydroxylysines.
 CC !- PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC !- SIMILARITY: Belongs to the fibrillar collagen family.
 DR PIR; A02B6; CG075;
 DR InterPro; IPR008161; Clg helix.
 DR InterPro; IPR008160; Collagen.
 DR PF01391; Collagen; 17.
 DR ProDom; PD000007; Clg helix; 4.
 DR PDB; 1A208; VWF_C1; PART1.
 KW Collagen; Direct protein sequencing; Extracellular matrix; Glycoprotein; Hydroxylation; Repeat; Structural protein.
 FT DOMAIN 1 14 Nonhelical region (N-terminal).
 FT DOMAIN 15 1040 Nonhelical region (C-terminal).
 FT DOMAIN 1041 1049 5-hydroxylysine.
 FT MOD_RES 95 95 5-hydroxylysine.
 FT MOD_RES 107 107 5-hydroxylysine.
 FT MOD_RES 119 119 5-hydroxylysine.
 FT MOD_RES 938 938 5-hydroxylysine.
 FT MOD_RES 950 950 5-hydroxylysine.
 FT CARBOHYD 107 107 O-linked (Gal. .).
 FT DISULFID 950 950 Interchain.
 FT DISULFID 1040 1040 Interchain.
 FT DISULFID 1041 1041 Interchain.
 SQ SEQUENCE 1049 AA; 93651 MW; 88EC33D1C66EC9A3 CRC64;

Query Match 40.8%; Score 60; DB 1; Length 1049;
 Best Local Similarity 57.1%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

447 GPQGPAGKNGETGPQGPPGFT 467

Db 076271 PRELIMINARY; PRT; 904 AA.
 ID 076271; DT 01-NOV-1998 (TREMBLrel. 08, Created)
 AC DT 01-OCT-2003 (TREMBLrel. 08, Last sequence update)
 RX DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 RA Qin X.X., Waite J.H.;
 RT "A potential mediator of collagenous block copolymer gradients in mussel byssal threads." ;
 RT RL Proc. Natl. Acad. Sci. U.S.A. 95:10517-10522(1998).
 RX DR EMBL; AF043944; AAC33847.1; -
 DR GO:0005737; C:cytoplasm; IEA.
 DR GO:0006817; P:phosphate transport; IEA.
 DR InterPro; IPR008161; Clg_helix.
 DR InteraPro; IPR008160; Collagen; 7.
 DR Pfam; PF01391; Collagen; 7.
 DR ProDom; PD000007; Clg_helix; 1.
 SQ SEQUENCE 904 AA; 77883 MW; 5529135651AD4C40 CRC64;
 Query Match 40.5%; Score 59.5; DB 2; Length 904;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02; Indels 1; Gaps 1;
 Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 1 GMQGPAGKNGETGPQGPPGFTPLFSP 26
 DB 227 GPRGPAGKNGETGPQGPPG-PPGHS 251

RESULT 11

QBMW55 PRELIMINARY; PRT; 905 AA.
 ID QBMW55 AC QBMW55; DT 01-OCT-2002 (TREMBLrel. 22, Created)
 RX DR EMBL; AF043945; AAC33848.1; -
 DR GO:0005737; C:cytoplasm; IEA.
 DR InterPro; IPR008160; Collagen; 7.
 DR Pfam; PF01391; Collagen; 7.
 DR ProDom; PD000007; Clg_helix; 1.
 SQ SEQUENCE 904 AA; 77883 MW; 5529135651AD4C40 CRC64;
 Query Match 40.5%; Score 59.5; DB 2; Length 904;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02; Indels 1; Gaps 1;

QY 1 GMQGPAGKNGETGPQGPPGFTPLFSP 26
 DB 227 GPRGPAGKNGETGPQGPPG-PPGHS 251

RESULT 12

QBMW55 PRELIMINARY; PRT; 905 AA.
 ID QBMW55 AC QBMW55; DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 RX DR EMBL; AF043945; AAC33848.1; -
 DR GO:0005737; C:cytoplasm; IEA.
 DR InterPro; IPR008160; Collagen; 7.
 DR Pfam; PF01391; Collagen; 7.
 DR ProDom; PD000007; Clg_helix; 1.
 SQ SEQUENCE 904 AA; 77883 MW; 5529135651AD4C40 CRC64;
 Query Match 40.5%; Score 59.5; DB 2; Length 905;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02; Indels 3; Mismatches 9; Indels 1; Gaps 1;

DR GO; GO:0005581; C:collagen; IEA.
 DR GO; GO:005737; C:cytoplasm; IEA.
 DR GO; GO:005201; F:extracellular matrix structural constituent; IEA.
 DR GO; GO:008817; P:phosphate transport; IEA.
 DR InterPro; IPR00816; Collagen.
 DR InterPro; IPR000885; Fib_collagen_C.
 DR InterPro; IPR009041; PMP_SGCC.
 DR InterPro; IPR010077; VWF_C.
 DR PFam; PF01410; COLIFI_1.
 DR PFam; PF01391; Collagen; 16.
 DR PFam; PF00093; VWC_1.
 DR PRODom; PDO02078; Fib_collagen_C.
 DR SMART; SM00038; COLIFI_1.
 DR SMART; SM00214; VWC_1.
 DR PROSITE; PS01208; VWFC_1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Collagen. SEQUENCE 1447 AA; 137144 MW; 9CABD561F5BA36BF CRC64;
 Query Match 40.1%; Score 59; DB 2; Length 1447;
 Best Local Similarity 46.2%; Pred. No. 2.4e+02;
 Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
 Qy 1 GNGPAGSSGWEBSGSPPGVTPLFSP 26
 Db 607 GPAGPAGERGBQAGGPFQSLPGP 632

RESULT 15
 Q95LJ6 PRELIMINARY; PRT; 135 AA.
 ID Q95LJ6
 AC Q95LJ6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Butheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Macaca.
 NCBI_TaxID:9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE:Testis;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Hirata M., Suto Y., Hirai M.,
 RA Terao K., Sugano S., Hashimoto K.;
 RT "Cynomolgus monkey testicular cDNAs for discovery of novel human genes
 RT in the human genome sequence.";
 RL BMC Genomics 3:36-36 (2002).
 DR EMBL; AB027292; BAB69761.1; -.
 RW HYPOTHETICAL protein.
 SQ SEQUENCE 135 AA; 14280 MW; 398CC2768FFA25D7 CRC64;
 Query Match 39.8%; Score 58.5; DB 2; Length 135;
 Best Local Similarity 45.5%; Pred. No. 24;
 Matches 15; Conservative 1; Mismatches 8; Indels 9; Gaps 2;
 Qy 3 QGPAG-SGWEEGSGP-----PGYTPFLFSP 26
 Db 82 RGPVGASGWAEAGCASPQTSITPYPRVTPCQVP 114

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 29, 2005, 01:01:09 ; Search time 41 Seconds
(without alignments)
47.338 Million cell updates/sec

Title: US-09-980-263-1

Perfect score: 147

Sequence: 1 GHQGPASSGWERGSGSPPGVTPFLSP 26

Scoring table: BLOSUM62

Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Issued Patents_AAI*

1: /cgn2_6/_prodatal/1/iaa/5A_COMB_pep:
2: /cgn2_6/_prodatal/1/iaa/5B_COMB_pep:
3: /cgn2_6/_prodatal/1/iaa/6A_COMB_pep:
4: /cgn2_6/_prodatal/1/iaa/6B_COMB_pep:
5: /cgn2_6/_prodatal/1/iaa/PTC5_COMB_pep:
6: /cgn2_6/_prodatal/1/iaa/backfiles1_pep:*

Pred. No. is the number of results predicted by chance to have a a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	26	2 US-09-370-156-27	Sequence 27, App1
2	147	100.0	31	2 US-08-370-156-11	Sequence 11, App1
3	147	100.0	31	3 US-08-990-065-23	Sequence 23, App1
4	147	100.0	31	3 US-08-975-084-3	Sequence 3, App1
5	147	100.0	31	4 US-09-332-13	Sequence 13, App1
6	147	100.0	600	2 US-08-370-056-4	Sequence 4, App1
7	147	100.0	600	3 US-08-14-095-4	Sequence 4, App1
8	147	100.0	600	3 US-08-975-084-1	Sequence 1, App1
9	140	95.2	25	3 US-08-975-084-2	Sequence 2, App1
10	60	40.8	54	1 US-07-972-032-59	Sequence 59, App1
11	60	40.8	54	1 US-08-642-255-71	Sequence 71, App1
12	60	40.8	93	1 US-07-972-032-60	Sequence 60, App1
13	60	40.8	93	1 US-07-972-032-61	Sequence 61, App1
14	60	40.8	633	1 US-08-642-255-73	Sequence 73, App1
15	60	40.8	1065	1 US-08-642-255-72	Sequence 72, App1
16	59	40.1	532	1 US-08-642-255-72	Sequence 9, App1
17	57.5	39.1	280	4 US-09-823-386-18	Sequence 18, App1
18	57.5	39.1	280	4 US-09-823-386-18	Sequence 18, App1
19	57	38.8	520	4 US-09-252-991A-29451	Sequence 29451, A
20	57	38.8	1024	3 US-08-931-091A-29451	Sequence 2, App1
21	57	38.8	1366	3 US-08-963-025-19	Sequence 19, App1
22	57	38.8	1366	3 US-09-500-811-19	Sequence 19, App1
23	57	38.8	1366	3 US-09-570-573-19	Sequence 19, App1
24	57	38.8	1366	3 US-09-148-008-19	Sequence 19, App1
25	57	38.8	1366	4 US-09-555-887-10	Sequence 10, App1
26	57	38.8	1366	4 US-09-289-578-10	Sequence 10, App1
27	57	38.8	1366	4 US-09-949-016-5882	Sequence 5882, App

ALIGNMENTS

RESULT 1
US-08-370-156-27
; Sequence 27, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION: ← 
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR NUMBER OF SEQUENCES: 27
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ettington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 40999
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohl, Kenneth I.
; REGISTRATION NUMBER: 30,955
; DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-370-156-27
; Query Match 100.0%; Score 147; DB 2; Length 26;
; Best Local Similarity 100.0%; Pred. No. 7.e-11;
; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
 US-08-370-156-11
 / Sequence 11, Application US/08370156
 / Patent No. 592780
 / GENERAL INFORMATION:
 / APPLICANT: Soreq, Hermona
 / APPLICANT: Zakut, Haim
 / APPLICANT: Shani, Moshe
 / TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
 / TITLE OF INVENTION: ANTOCHOLINESTERASE SUBSTANCES
 / NUMBER OF SEQUENCES: 27
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Reising, Ethington, Barnard & Perry
 / STREET: P.O. Box 4390
 / CITY: Troy
 / STATE: Michigan
 / COUNTRY: US
 / ZIP: 48099
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/990,065
 / FILING DATE:
 / CLASSIFICATION: 514
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/850,347
 / FILING DATE: 02-MAY-1997
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/318,826
 / FILING DATE: 01-JAN-1994
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Montgomery, Ilene N.
 / REGISTRATION NUMBER: 38,972
 / REFERENCE/DOCKET NUMBER: 2391.00086
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (248) 539-5050
 / TELEFAX: (248) 539-5055
 / INFORMATION FOR SEQ ID NO: 23:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 31 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: peptide
 / US-08-990-065-23

Query Match 100.0%; Score 147; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
 US-08-990-065-23
 / Sequence 23, Application US/08990065
 / Patent No. 6121046
 / GENERAL INFORMATION:
 / APPLICANT: Soreq, Hermona
 / APPLICANT: Seidman, Shlomo
 / APPLICANT: Eckstein, Fritz
 / APPLICANT: Friedman, Alon
 / APPLICANT: Kauter, Daniela
 / TITLE OF INVENTION: SYNTHETIC ANTISENSE
 / TITLE OF INVENTION: OLIGODEOXYNUCLEOTIDES AND PHARMACEUTICAL COMPOSITIONS
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Kohn & Associates
 / STREET: 30500 No. 6121046thwestern Hwy. Suite 410
 / CITY: Farmington Hills
 / STATE: Michigan
 / COUNTRY: U.S.
 / ZIP: 48334

Query Match 100.0%; Score 147; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
 US-08-975-084-3
 / Sequence 3, Application US/08975084
 / Patent No. 6258780
 / GENERAL INFORMATION:
 / APPLICANT: Soreq, Hermona
 / APPLICANT: Seidman, Shlomo
 / APPLICANT: Friedman, Alon
 / APPLICANT: Kauter, Daniela
 / TITLE OF INVENTION: A METHOD AND COMPOSITION FOR ENABLING
 / PASSAGE THROUGH THE BLOOD-BRAIN BARRIER
 / NUMBER OF SEQUENCES: 9
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Kohn & Associates
 / STREET: 30500 No. 6258780thwestern Hwy., Suite 410
 / CITY: Farmington Hills
 / STATE: Michigan
 / COUNTRY: US
 / ZIP: 48334

Query Match 100.0%; Score 147; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.2e-11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 GMQGPAGSGWEGSGSGSPGPVTPFLFSP 26
 6 GMQGPAGSGWEGSGSGSPGPVTPFLFSP 31

Db

TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055
 SEQUENCE CHARACTERISTICS:
 LENGTH: 31 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-975-084-3

Query Match Similarity 100.0%; Pred. No. 9.2e-11; Length 31;
 Best Local Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 26
 Db 6 GMQGPAGSGWEEGSGSPPGVTPFLFSP 31

RESULT 5
 US-09-380-532-13
 Sequence 13, Application US/09380532
 Patent No. 6475998

GENERAL INFORMATION:
 APPLICANT: Soreq, Hermona
 Seidman, Shlomo
 Shohami, Esther

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF INJURY TO THE CENTRAL NERVOUS SYSTEM
 NUMBER OF SEQUENCES: 13
 CURRENT APPLICATION DATA:
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kohn & Associates
 STREET: 30100 N. 6475998thwestern Hwy.
 CITY: Farmington Hills
 STATE: Michigan
 ZIP: 48334

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/380.532
 FILING DATE: 12-No. 6475998-1999
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Montgomery, Ilene N.
 REGISTRATION NUMBER: 38,972
 REFERENCE/DOCKET NUMBER: 2391.00089
 TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055

INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 31 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Query Match Similarity 100.0%; Pred. No. 9.2e-11; Length 31;
 Best Local Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 26
 Db 6 GMQGPAGSGWEEGSGSPPGVTPFLFSP 31

RESULT 6
 US-08-370-156-4
 Sequence 4, Application US/08370156

GENERAL INFORMATION:
 APPLICANT: Soreq, Hermona
 ZAKUT, Haim
 APPLICANT: Shani, Moshe
 TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Reising, Ettington, Barnard & Perry
 STREET: P.O. Box 4390
 CITY: Troy
 STATE: Michigan
 COUNTRY: US
 ZIP: 48099

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/370.156
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Kohn, Kenneth I.
 REGISTRATION NUMBER: 10, 955
 REFERENCE/DOCKET NUMBER: P-307 (Mulford)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (810) 689-3500
 TELEFAX: (810) 689-0711
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 600 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-370-156-4

Query Match Similarity 100.0%; Pred. No. 1.8e-09;
 Best Local Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 26
 Db 575 GMQGPAGSGWEEGSGSPPGVTPFLFSP 600

RESULT 7
 US-08-814-095-4
 Sequence 4, Application US/08814095

GENERAL INFORMATION:
 Patent No. 6025183

APPLICANT: Soreq, Hermona
 ZAKUT, Haim
 APPLICANT: Shani, Moshe
 TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR NUMBER OF SEQUENCES: 7
 TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: KOHN & ASSOCIATES
 STREET: 30500 No. 6025183thwestern Highway, Suite 410
 CITY: Farmington Hills
 STATE: Michigan
 COUNTRY: U.S.

Query Match Similarity 100.0%; Pred. No. 9.2e-11; Length 31;
 Best Local Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMQGPAGSGWEEGSGSPPGVTPFLFSP 26
 Db 6 GMQGPAGSGWEEGSGSPPGVTPFLFSP 31

ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/814,095
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Montgomery, Ilene N.
 REGISTRATION NUMBER: 38,972
 REFERENCE/DOCKET NUMBER: 2391.00066
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055
 SEQUENCE CHARACTERISTICS:
 LENGTH: 600 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-814-095-4

Query Match 100.0%; Score 147; DB 3; Length 600;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 147; DB 3; Length 600;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
 US-08-975-084-2
 Sequence 2, Application US/08975084
 Patent No. 6258750
 GENERAL INFORMATION:
 APPLICANT: SOREQ, Hermona
 APPLICANT: FRIEDMAN, Alon
 APPLICANT: SEIDMAN, Shlomo
 APPLICANT: KAUFER, Daniela
 TITLE OF INVENTION: A METHOD AND COMPOSITION FOR ENABLING
 TITLE OF INVENTION: PASSAGE THROUGH THE BLOOD-BRAIN-BARRIER
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kohn & Associates
 STREET: 3050 No. 6258780thwestern Hwy., Suite 410
 CITY: Farmington Hills
 STATE: Michigan
 COUNTRY: US
 ZIP: 48334
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/975,084
 FILING DATE: 11-NOV-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Montgomery, Ilene N.
 REGISTRATION NUMBER: 38,972
 REFERENCE/DOCKET NUMBER: 2391.00082
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (248) 539-5050
 TELEFAX: (248) 539-5055
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO

US-08-975-084-2

Query Match 95.2%; Score 140; DB 3; Length 25;
 Best Local Similarity 100.0%; Pred. No. 4.8e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 147; DB 3; Length 600;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 575 GMQGPAGSGWEGSSPPGVTPLFS 600

RESULT 10
 US-07-972-032-59

RESULT 14
 US-08-642-255-73
 Sequence 73, Application US/08642255
 ; Patent No. 5773249
 ; GENERAL INFORMATION:
 ; APPLICANT: CAPPELLO, Joseph
 ; FERRARI, Franco A.
 ; TITLE OF INVENTION: High Molecular Weight Collagen-Like
 ; TITLE OF INVENTION: Protein Polymers
 ; NUMBER OF SEQUENCES: 135
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOHACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-1187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/642,255
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ROWLAND, Bertram I.
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: A55555-3/BIR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 494-8700
 ; TELEFAX: (415) 494-8771
 ; TELEX: 910 277299 FHT UR
 ; INFORMATION FOR SEQ ID NO: 73:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 633 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-642-255-73

Query Match 1 GMQGPAGSGWEGSGSPPG 19
 Db 52 GAQGPAGPGGSRGDPGPPG 70

RESULT 15
 US-08-642-255-72
 Sequence 72, Application US/08642255
 ; Patent No. 5773249
 ; GENERAL INFORMATION:
 ; APPLICANT: CAPPELLO, Joseph
 ; FERRARI, Franco A.
 ; TITLE OF INVENTION: High Molecular Weight Collagen-Like
 ; TITLE OF INVENTION: Protein Polymers
 ; NUMBER OF SEQUENCES: 135
 ; CORRESPONDENCE ADDRESS:
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-1187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

Query Match 1 GMQGPAGSGWEGSGSPPG 19
 Db 43 GAQGPAGPGGSRGDPGPPG 61

RESULT 13
 US-07-972-032-61
 Sequence 61, Application US/07972032
 ; Patent No. 5496712
 ; GENERAL INFORMATION:
 ; APPLICANT: Cappello, Joseph
 ; FERRARI, Franco A.
 ; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE
 ; TITLE OF INVENTION: PROTEIN POLYMERS
 ; NUMBER OF SEQUENCES: 85
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bertram I. Rowland
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: CA 94111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/972,032
 ; FILING DATE: 1992-11-05
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/791,960
 ; FILING DATE: 12-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Bertram I.
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: A-55555-1/BIR; PROP-08-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 398-1989
 ; TELEFAX: (415) 398-3249
 ; INFORMATION FOR SEQ ID NO: 61:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 93 amino acids
 ; TYPE: AMINO ACID
 ; MOLECULE TYPE: protein
 ; FEATURE: linear
 ; NAME/KEY: Duplication
 ; LOCATION: 34 .. 69
 ; OTHER INFORMATION: /label= internalduplica
 ; OTHER INFORMATION: /note= "34-42 x 2; 43-60 x 6; 61-69 x 2; 34-69 x
 ; OTHER INFORMATION: 4."
 ; US-07-972-032-61

Query Match 1 GMQGPAGSGWEGSGSPPG 19
 Db 43 GAQGPAGPGGSRGDPGPPG 61

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROWLAND, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A55556-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1065 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-255-72

Query Match 40.8%; Score 60; DB 1; Length 1065;
Best Local Similarity 57.9%; Pred. No. 44;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 GMQGPAGSGWEEGSGSPRG 19
Db 52 GAQGPAGPGSGRDPGPPG 70

Search completed: October 29, 2005, 01:05:21
Jqb time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 01:03:40 ; Search time 165 Seconds
(without alignments)
65.873 Million cell updates/sec

Title: US-09-980-263-1

Perfect score: 147
Sequence: 1 GRQGPASSGWEBSGSPPGVTPLFSSP 26

Scoring table: BLOSUM62
Gappp 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : Published Applications AA:*

- /cgn2_6/_ptodata/1/_pubpaas/US07_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/PCT_NEW_PUB.pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US06_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US05_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US04_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US03_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US02_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US09_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US05C_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US09_NEW_PUB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10A_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10B_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10C_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10D_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10E_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US10F_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US11A_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US11B_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US11C_PUBCOMB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US60_NEW_PUB_pep:*
- /cgn2_6/_ptodata/1/_pubpaas/US60_PUBCOMB_pep:*

RESULT 1
US-09-980-042-1
; Sequence 1, Application US/09998042
; Publication No. US2003003632A1
; GENERAL INFORMATION:
; APPLICANT: VISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW
; TITLE OF INVENTION: ACETYLCHOLINESTERASE-DERIVED PEPTIDE AND USES THEREOF
; FILE REFERENCE: 7811/WO/99
; CURRENT APPLICATION NUMBER: US/09/998, 042
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-980-042-1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	26	US-09-980-042-1	Sequence 1, Appli
2	147	100.0	37	US-09-980-042-6	Sequence 6, Appli
3	147	100.0	53	US-09-980-042-7	Sequence 7, Appli
4	68	46.3	105	US-10-425-115-356384	Sequence 156384,
5	61.5	41.8	157	US-10-425-115-274774	Sequence 274774,
6	61.5	41.8	575	US-10-276-774-1845	Sequence 1845, Ap
7	61	41.5	138	US-10-425-115-22126	Sequence 22126,
8	60	40.8	334	US-10-420-155-131	Sequence 131, App
9	60	40.8	334	US-10-820-155-133	Sequence 250621,
10	60	40.8	528	US-10-425-115-250621	Sequence 70635, A
11	60	40.8	534	US-10-425-114-70635	

Qy	1	GMQGPAGSGWEGSGSPPGVTPFLFSP	26
Db	1	GMQGPAGSGWEGSGSPPGVTPFLFSP	26
RESULT 2 US-09-980-042-6			
; Sequence 6, Application US/09998042			
; Publication No. US2003003632A1			
; GENERAL INFORMATION:			
; APPLICANT: VISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW			
; TITLE OF INVENTION: ACETYLCHOLINESTERASE-DERIVED PEPTIDE AND USES THEREOF			
; FILE REFERENCE: 7811/WO/99			
; CURRENT APPLICATION NUMBER: US/09/998, 042			
; CURRENT FILING DATE: 2003-07-02			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 26			
; TYPE: PRT			
; ORGANISM: HOMO SAPIENS			
US-09-980-042-1			

FILE REFERENCE: 7811/NO/99
 CURRENT APPLICATION NUMBER: US/09/998,042
 CURRENT FILING DATE: 2002-07-02
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 6
 LENGTH: 37
 TYPE: PRT
 ORGANISM: homo sapiens
 US-09-998-042-6

Query Match 100.0%; Score 147; DB 10; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1..4e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 US-10-425-115-274774
 Query Match 46.3%; Score 68; DB 16; Length 105;
 Best Local Similarity 61.1%; Pred. No. 3..3;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GMQPAGSGNNEEGSGSPPGTTPLFSP 18
 Db 80 GTRGGGSGWPSFGNPP 97

Query Match 100.0%; Score 147; DB 10; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1..4e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 US-10-425-115-274774
 Query Match 46.3%; Score 68; DB 16; Length 105;
 Best Local Similarity 61.1%; Pred. No. 3..3;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GMQPAGSGNNEEGSGSPPGTTPLFSP 26
 Db 12 GMQPAGSGNNEEGSGSPPGTTPLFSP 37

RESULT 3
 US-09-998-042-7
 Sequence 7, Application US/0998042
 Publication No. US2003036632A1
 GENERAL INFORMATION:
 APPLICANT: VISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW
 TITLE OF INVENTION: ACETYLCHOLINESTERASE-DERIVED PEPTIDE AND USES THEREOF
 FILE REFERENCE: 7811/NO/99
 CURRENT APPLICATION NUMBER: US/09/998,042
 CURRENT FILING DATE: 2002-07-02
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 7
 LENGTH: 53
 TYPE: PRT
 FEATURE: OTHER INFORMATION: Description of Artificial Sequence:arp-two hybrid
 OTHER INFORMATION: Screen peptide
 US-09-998-042-7

Query Match 100.0%; Score 147; DB 10; Length 53;
 Best Local Similarity 100.0%; Pred. No. 2e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 US-10-425-115-274774
 Query Match 46.3%; Score 68; DB 16; Length 105;
 Best Local Similarity 61.1%; Pred. No. 3..3;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GMQPAGSGNNEEGSGSPPGTTPLFSP 26
 Db 28 GMQPAGSGNNEEGSGSPPGTTPLFSP 53

RESULT 6
 US-10-426-774-1845
 Sequence 1845, Application US/10276774
 Publication No. US20040053245A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq, Inc.
 APPLICANT: Tang, Y, Tom et al
 APPLICANT: No. Tom et al
 TITLE OF INVENTION: Nucleic Acids and Polypeptides
 FILE REFERENCE: 21272-030
 CURRENT APPLICATION NUMBER: US/10/276,774
 CURRENT FILING DATE: 2002-11-18
 PRIOR APPLICATION NUMBER: 09/560,875
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: 2000-02-03
 NUMBER OF SEQ ID NOS: 2700
 SOFTWARE: Custom
 SEQ ID NO 1845
 LENGTH: 575
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(575)

RESULT 4
 US-10-425-115-356384
 Sequence 356384, Application US/10425115
 Publication No. US2004214272A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovacic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 FILE REFERENCE: 38-21(53222)B
 CURRENT APPLICATION NUMBER: US/10/425,115
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 369326
 SEQ ID NO 356384
 LENGTH: 105
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(105)
 OTHER INFORMATION: unsure at all Xaa locations

OTHER INFORMATION: Xaa = any amino acid or nothing

us-10-276-774-1845

Query Match Score 41.8%; Best Local Similarity 56.5%; Pred. No. 88; Matches 13; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Qy 1 GMQG-PAGSGWEGSGSPGVTP 22

Db 421 GQGQFLFGWGLKEGCLPPGPP 443

RESULT 7

US-10-425-115-221226

Sequence 221226, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Zhou, Yinhua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TYPE: PRT

ORGANISM: Zea mays

FEATURE: NAME/KEY: unsure

LOCATION: (1) (138)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE: OTHER INFORMATION: Clone ID: MRT4577_133350C.1.pep

US-10-425-115-221226

Query Match Score 41.5%; Best Local Similarity 52.4%; Pred. No. 26; Matches 11; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

Qy 4 GPAGS--GWECSGSPPGVTP 22

Db 95 GPMGTGLWWRPGAGGPPGPPIP 115

RESULT 8

US-10-820-155-131

Sequence 131, Application US/10820155

Publication No. US20050137126A1

GENERAL INFORMATION:

APPLICANT: Natimmune A/S

APPLICANT: Weilgony, Dietmar

APPLICANT: Jensenius, Jens Christian

APPLICANT: Kongerslev, Leif

APPLICANT: Matthiesen, Finn

TITLE OF INVENTION: Treatment of SARS in individuals

FILE REFERENCE: P-774 US00

CURRENT APPLICATION NUMBER: US/10/820,155

CURRENT FILING DATE: 2004-04-08

NUMBER OF SEQ ID NOS: 137

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 131

LENGTH: 334

TYPE: PRT

ORGANISM: Mus musculus

OTHER INFORMATION: Other Molecules Associated With

US-10-820-155-133

Query Match Score 60; Best Local Similarity 50.0%; Pred. No. 1.2e+02; Matches 14; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

Qy 1 GMQGPA--GSGWEGSGSPGVTPFLFSP 26

Db 7 GRAGAIKGRGWKIGSFVDFGVFPVLS 34

RESULT 11

US-10-425-114-70635

Query Match Score 40.8%; Best Local Similarity 60.0%; Pred. No. 78; Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21153311.B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO: 70635
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFBLB73016D03_FLI_Pep
US-10-425-114-7-0635

RESULT 12
Query Match 40.8%; Score 60; DB 15; Length 534;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 14; Conservative 2; Mismatches 10; Indels 2; Gaps 1;
SEQ ID NO: 4
Qy 1 GMQGPAA-GSSWEEGSSGSPGPPGTYPLFSP 26
Db 13 GRAGGAIKGRWKGXGSGFVDDSVFPVLSP 40

RESULT 13
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

RESULT 14
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

RESULT 15
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

RESULT 16
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

RESULT 17
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

RESULT 18
Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
SEQ ID NO: 6
Qy 1 GMQGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPSPGPT 623

```

ORGANISM: *Sus scrofa*
us-10-402-072A-6
Query Match Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 GMQPAGSGWBEGSPPGVT 21
Db 603 GPQGPAGRNNGTGPQGPPT 623

Search completed: October 29, 2005, 01:18:08
Job time : 166 secs